

**Table S1.** Description of the strains included in this work indicating their origin, source, the total number of valid CDSs and their application (set). Strains labeled as “Hetero” or “Small” were analyzed but eventually discarded from the pipeline analysis.

Strain	Set	# CDSs	Source	Origin	Description	Ref
273614N	Learning	6148	CLIN	UK	Isolated from feces at the Royal Victoria Infirmary, Newcastle UK.	Liti et al. 2009
322134S	Learning	6363	CLIN	UK	Isolated from sputum at the Royal Victoria Infirmary, Newcastle UK.	Liti et al. 2009
378604X	Learning	6156	CLIN	UK	Isolated from sputum at the Royal Victoria Infirmary, Newcastle UK.	Liti et al. 2009
BC187	Learning	6368	WINE	California, USA	Found in barrels for fermentation in the Napa Valley	Liti et al. 2009
DBVPG1106	Learning	6331	WINE	Australia	Isolated from grapes	Liti et al. 2009
DBVPG1373	Learning	6382	SOIL	Netherlands	Isolated from soil	Liti et al. 2009
DBVPG1788	Learning	6357	SOIL	Finland	Isolated from soil	Liti et al. 2009
DBVPG1853	Learning	6324	BAKER	Ethiopia	Used to prepare the BAKER with white teff, a very small no glutine cereal typical of Ethiopia	Liti et al. 2009
DBVPG6040	Learning	6342	FRUIT	Netherlands	(aka <i>S. fructum</i> )	Liti et al. 2009
DBVPG6044	Learning	6389	WINE	West Africa	a.k.a. <i>S. mangini</i> , used to produce the Bili wine, from the tubercles of <i>Osbeckia grandiflora</i>	Liti et al. 2009
DBVPG6765	Learning	6427	WINE	Indonesia	a.k.a. <i>S. bouldarii</i> and NYCY3264) its habitat is the Lichee fruit and can be used to make a Lechee wine (typical chinese recipe)	Liti et al. 2009
K11	Learning	6356	SAKE	Japan	Shochu sake strain	Liti et al. 2009
L_1374	Learning	6384	WINE	Chile	Use in fermentation from must Cabernet	Liti et al. 2009
L_1528	Learning	6380	WINE	Chile	Fermentation from must Pais	Liti et al. 2009
NCYC110	Learning	6334	BEER	West Africa	a.k.a. <i>S. chevalierii</i> , used to produce the Ginger beer from <i>Z. officinalis</i>	Liti et al. 2009
NCYC361	Learning	6326	BEER	Ireland	a.k.a. <i>Diastaticus</i> , found in spoiled wort (malt must) during beer production	Liti et al. 2009
REF	Learning	6421	LAB		Is a resequencing of thre reference genome S288C	Liti et al. 2009
RM11_1A	Learning	6533	WINE	California, USA	Isolated in a vinjard	Liti et al. 2009
S288c	Learning	6418	LAB		Refrence strain	SGD
SK1	Learning	6285	SOIL	Illinois, USA	Laboratory strain used as a model for meiosis	Liti et al. 2009
UWOPS03_46 1_4	Learning	6339	FRUIT	Malaysia	Nectar, Bertram palm	Liti et al. 2009
UWOPS05_21 7_3	Learning	6282	FRUIT	Malaysia	Nectar, Bertram palm	Liti et al. 2009
UWOPS05_22 7_2	Learning	6368	FRUIT	Malaysia	Trigona sp. near Bertram palm	Liti et al. 2009

UWOPS83_78_7_3	Learning	6372	FRUIT	Bahamas	Fruit, <i>Opuntia stricta</i>	Liti et al. 2009
UWOPS87_24_21	Learning	6496	FRUIT	Hawaii	Cladode <i>Opuntia megacantha</i> near Bertram palm	Liti et al. 2009
W303	Learning	6294	LAB		Laboratory strain, a bud4 mutation causes haploids to bud with a mixture of axial and bipolar patterns	yeastgenome.org
Y12	Learning	6414	WINE	Ivory Coast, Africa	Isolated from palm wine in Ivory Coast, Africa, pre-1981.	genome.wustl.edu
Y55	Learning	6336	WINE	France	Laboratory strain a.k.a. 52	Liti et al. 2009
Y9	Learning	6139	SAKE	Java, Indonesia	Isolated from Ragi (African or finger millet, a sake like wine) in Java, Indonesia, pre-1962.	genome.wustl.edu
Yllc17_E5	Learning	6467	WINE	Sauternes, France		Liti et al. 2009
YJM789	Learning	6359	CLIN	USA	Isolated from the lung of an AIDS patient with pneumonia	med.stanford.edu
YJM975	Learning	6168	CLIN	Bergamo, Italy	Ospedali Riuniti di Bergamo, Italy, 1994-6, Isolated from vagina of patient suffering from vaginitis	Liti et al. 2009
YJM978	Learning	6372	CLIN	Bergamo, Italy	Ospedali Riuniti di Bergamo, Italy, 1994-6, Isolated from vagina of patient suffering from vaginitis	Liti et al. 2009
YJM981	Learning	6303	CLIN	Bergamo, Italy	Ospedali Riuniti di Bergamo, Italy, 1994-6, Isolated from vagina of patient suffering from vaginitis	Liti et al. 2009
YPS128	Learning	6352	SOIL	Pennsylvania, USA	Isolated from soil in a woodland, close to a <i>Quercus alba</i> (oak)	Liti et al. 2009
YPS606	Learning	6360	OAK		Isolated in a woodland, in the bark of a <i>Quercus robur</i> (oak). Spore derivative of YPS142	Liti et al. 2009
YS2	Learning	6336	BAKER	Australia	Baker strain	Liti et al. 2009
YS4	Learning	6136	BAKER	Netherlands	Baker strain	Liti et al. 2009
YS9	Learning		BAKER	Singapore	Lesaffre commercial baker strain	Liti et al. 2009
EC1118	Validation	6349	WINE	Champagne, France	Isolated, studied and selected from Champagne fermentations. Due to its competitive factor and ability to ferment equally well over a wide temperature range, the EC-1118 is one of the most widely used yeasts in the world.	Lalvin
JAY291	Validation	6140	ETOH	Brazil	Sugar cane fermentation for the production of bioethanol, a haploid derivative of PE-2	Argueso et al. 2009
RM11_1A	Validation	6282	WINE	California, USA	Isolated in a vinjard	Liti et al. 2009
Sigma1278b	Validation	6416	LAB		Laboratory strains used in pseudohyphal growth studies	yeastgenome.org
T7	Validation	6295	OAK	Missouri, USA	Isolated from oak tree exudates in Babler State Park, Missouri in 2003.	genome.wustl.edu
UC5	Validation	5598	SAKE	Japan	<i>S. cerevisiae</i> strain UC5 (UCD 612) was isolated from Sene sake	genome.wustl.edu

					in Kurashi, Japan, pre-1974.	
W303	Validation	6437	LAB		Laboratory strain, contains a bud4 mutation that causes haploids to bud with a mixture of axial and bipolar budding patterns	yeastgenome.org
YJM789	Validation	6374	CLIN	USA	Isolated from the lung of an AIDS patient with pneumonia	med.stanford.edu
AWRI796	Hetero	5568	WINE	South Africa	Isolated from South Africa, this strain is used extensively for its ability to allow the full expression of varietal fruit and good palate structure.	Borneman et al. 2011
FostersB	Hetero	2064	BEER		Industrial strains used in the Fosters breweries	Borneman et al. 2011
FostersO	Hetero	2735	BEER		Industrial strains used in the Fosters breweries	Borneman et al. 2011
LalvinQA23	Hetero	2911	WINE	Portugal	Selected in Portugal by the University of Trás-os-Montes e Alto Douro (UTAD) in cooperation with the Viticultural Commission of the Region Vinhos Verdes. In France QA23 was first introduced to produce wines for the English and German markets.	Borneman et al. 2011
Vin13	Hetero	3067	WINE	South Africa	Product of the yeast hybridization program of the Institute for Wine Biotechnology at the University of Stellenbosch in South Africa.	ScottLab
VL3	Hetero	4041	WINE	Bordeaux, France	VL3 has a special capability for revealing Sauvignon Blanc varietal aroma, by reactions involving the odorless precursors present in the must, particularly S-cysteine conjugates.	Winemaking jackkeller.net
AWRI1631	Small	4738	WINE	South Africa	A haploid derivative of the commercial wine strain N96 with identical fermenting capability of the N96	Borneman et al. 2011
CBS7960	Small	5332	ETOH	Sao Paulo, Brazil	Was isolated from in a factory in Sao Paulo, Brazil, which produces ethanol from cane-sugar syrup.	genome.wustl.edu
CLIB215	Small	5049	BAKER	New Zealand	A baker's yeast from New Zealand collected in 1994.	genome.wustl.edu
CLIB324	Small	3732	BAKER	Saigon, Vietnam	A baker's yeast used in Saigon, Vietnam and collected in 1996.	genome.wustl.edu
CLIB382	Small	937	BEER	Ireland	Isolated from super-attenuated beer prepared in an Irish brewery pre-1952.	genome.wustl.edu
FL100	Small	3908	LAB		S. cerevisiae strain FL100 is a laboratory strain collected pre-1968.	genome.wustl.edu
M22	Small	2547	WINE	Italy	A strain isolated from a vineyard in Italy	Doniger et al. 2008
PW5	Small	5522	WINE	Abia, Nigeria	Isolated from Raphia palm wine in Aba, Abia state, Nigeria in 2002.	genome.wustl.edu
T73	Small	3324	WINE	Alicante, Spain	Isolated from a red wine prepared	genome.wustl.edu

					from a Monastrel grape in fermentation stage in Alicante, Spain in 1987.	
Y10	Small	2655	FRUIT	Philippines	Isolated from a coconut in the Phillipines, pre-1973.	genome.wustl.edu
YJM269	Small	5395	WINE	Austria	Isolated from Blauer Portugieser grapes in 1954.	genome.wustl.edu
YPS163	Small	3151	OAK	USA	Isolated from an oak tree in the United States of America	Doniger et al. 2008

**Table S2.** Clusters of strains used to explore the genome wide phylogenetic relationships according to different definitions.

Name	# clusters	Categories
<b>Liti</b>	5	Reference / West African / Malaysian / North American / Wine-European
<b>Geo1</b>	5	Usa+bahamas / Oceania / Europa / Africa / South America / Reference
<b>Geo2</b>	10	South America / USA / Bahamas / Australia / Indonesia_Malaysia / Japan / Hawaii / Africa / Europa / Reference
<b>Geo3</b>	14	Chile / Netherlands / Italy / Irland_UK / France / Finland / Eastern USA / California USA / Hawaii USA / Bahamas / Australia / Malaysia_Singapore_Indonesia / Japan / West Africa / Etiopia / Reference
<b>Geo4</b>	4	Oceania / Europa / South America / Reference
<b>Geo5</b>	4	Usa+bahamas / Europa / South America / Reference
<b>Source1</b>	6	Baker / Beer / Clinical from vagina / Clinical sputum / Wine fermentation / Palm / Fruit / Soil (optional/remove) / Rago_Sake
<b>Source2</b>	9	Baker_Beer / Clinical / Wine / Ref / Fruit / Rago_Sake / Soil

Liti	Geo1	Geo2	Geo3	Geo4	Geo5	Source1	Source2
REF	YPS606	L_1528	L_1528	YS2	YPS606	YS4	YS4
S288c	YPS128	L_1374	L_1374	DBVPG1106	YPS128	YS2	YS2
W303	UWOPS87_2421			YS9	UWOPS87_2421	YS9	YS9
	YJM789	YPS606	DBVPG6040	K11	YJM789	DBVPG1853	BC187
DBVPG6044	RM11_1A	YPS128	YS4	UWOPS03_461_4	RM11_1A	DBVPG6044	NCYC361
NCYC110	SK1	YJM789	DBVPG1373	UWOPS05_217_3	SK1	BC187	DBVPG1853
	BC187	RM11_1A		UWOPS05_227_2	BC187	NCYC361	
UWOPS03_461_4	UWOPS83_787_3	SK1	YJM978	Y9	UWOPS83_787_3	UWOPS03_461_4	YJM978
UWOPS05_217_3		BC187	YJM981			YJM978	YJM981
UWOPS05_227_2	YS2		YJM975	NCYC361	NCYC361	YJM981	YJM975
	DBVPG1106	YS2		DBVPG6040	DBVPG6040	YJM975	273614N
YPS128	YS9	DBVPG1106	NCYC361	YS4	YS4	YPS128	378604X
YPS606	K11		273614N	DBVPG1373	DBVPG1373	378604X	322134S
	UWOPS03_461_4	UWOPS03_461_4	378604X	YJM978	YJM978	322134S	YJM789
BC187	UWOPS05_217_3	UWOPS05_217_3	322134S	YJM981	YJM981	BC187	
DBVPG1106	UWOPS05_227_2	UWOPS05_227_2		YJM975	YJM975	Y55	YIle17_E5
DBVPG1373	Y9	Y9	YIle17_E5	273614N	273614N	DBVPG1106	L_1528
YJM975		YS9	Y55	378604X	378604X	YJM975	L_1374
YJM978	NCYC361			322134S	322134S	RM11_1A	Y55
DBVPG1788	DBVPG6040	UWOPS87_2421	YPS606	YIle17_E5	YIle17_E5	DBVPG6044	DBVPG1106
RM11_1A	YS4		YPS128	Y55	Y55	YIle17_E5	RM11_1A
L_1374	DBVPG1373	DBVPG6044	YJM789	DBVPG1788	DBVPG1788	L_1528	DBVPG6044
DBVPG6765	YJM978	NCYC110	SK1	DBVPG6765	DBVPG6765	L_1374	
L_1528	YJM981	DBVPG1853				L_1528	S288c
YJM981	YJM975	Y12	RM11_1A	L_1528	L_1528	UWOPS03_461_4	W303
	273614N		BC187	L_1374	L_1374	UWOPS05_217_3	
	378604X	NCYC361				Y12	DBVPG6040
	322134S	DBVPG6040	YS2	S288c	S288c		UWOPS83_787_3
	YIle17_E5	YS4	DBVPG1106	W303	W303	DBVPG6040	UWOPS03_461_4
	Y55	DBVPG1373				UWOPS83_787_3	UWOPS05_217_3
DBVPG1788	YJM978	YJM978	UWOPS03_461_4			UWOPS87_2421	Y12
DBVPG6765	YJM981	YJM981	UWOPS05_217_3				UWOPS87_2421
	YJM975	YJM975	UWOPS05_227_2				
	DBVPG6044	273614N	YS9			DBVPG1373	
	NCYC110	378604X	Y9			DBVPG1788	Y9
	DBVPG1853	322134S				SK1	K11
	Y12	YIle17_E5	DBVPG6044			YPS128	
		Y55	NCYC110				DBVPG1373
	L_1528	DBVPG1788	Y12			Y9	DBVPG1788
	L_1374	DBVPG6765				K11	SK1
			W303				YPS128
	S288c	S288c	S288c				
	W303	W303					

**Table S3.** Description of the 41 genes that presented no variability after alignment in all the 39 strains analyzed. All the information have been taken from the Saccharomyces Genome Database (SGD). Asterisks marks genes that were conserved in the 47 strains (39 from the Liti's learning set plus 8 from the validation set)

ORF	Name	Status	Chr	Description
*YBL029C-A		Unchar.	2	Protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cell periphery; has potential orthologs in Saccharomyces species and in Yarrowia lipolytica
*YBL092W	<i>RPL32</i>	Verified	2	Protein component of the large (60S) ribosomal subunit, has similarity to rat L32 ribosomal protein; overexpression disrupts telomeric silencing
*YDL061C	<i>RPS29B</i>	Verified	4	Protein component of the small (40S) ribosomal subunit; nearly identical to Rps29Ap and has similarity to rat S29 and E. coli S14 ribosomal proteins
*YDL133C-A	<i>RPL41B</i>	Verified	4	Ribosomal protein L47 of the large (60S) ribosomal subunit, identical to Rpl41Ap and has similarity to rat L41 ribosomal protein; comprised of only 25 amino acids; rpl41a rpl41b double null mutant is viable
YDL184C	<i>RPL41A</i>	Verified	4	Ribosomal protein L47 of the large (60S) ribosomal subunit, identical to Rpl41Bp and has similarity to rat L41 ribosomal protein; comprised of only 25 amino acids; rpl41a rpl41b double null mutant is viable
*YDL232W	<i>OST4</i>	Verified	4	Subunit of the oligosaccharyltransferase complex of the ER lumen, which catalyzes protein asparagine-linked glycosylation; type I membrane protein required for incorporation of Ost3p or Ost6p into the OST complex
YDR154C		Dubious	4	Dubious open reading frame, null mutant exhibits synthetic phenotype with alpha-synuclein
YDR155C	<i>CPR1</i>	Verified	4	Cytoplasmic peptidyl-prolyl cis-trans isomerase (cyclophilin), catalyzes the cis-trans isomerization of peptide bonds N-terminal to proline residues; binds the drug cyclosporin A
YER009W	<i>NTF2</i>	Verified	5	Nuclear envelope protein, interacts with GDP-bound Gsp1p and with proteins of the nuclear pore to transport Gsp1p into the nucleus where it is an essential player in nucleocytoplasmic transport
*YER044C	<i>ERG28</i>	Verified	5	Endoplasmic reticulum membrane protein, may facilitate protein-protein interactions between the Erg26p dehydrogenase and the Erg27p 3-ketoreductase and/or tether these enzymes to the ER, also interacts with Erg6p
*YER053C-A		Unchar.	5	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the endoplasmic reticulum
*YER056C-A	<i>RPL34A</i>	Verified	5	Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl34Bp and has similarity to rat L34 ribosomal protein
YGL030W	<i>RPL30</i>	Verified	7	Protein component of the large (60S) ribosomal subunit, has similarity to rat L30 ribosomal protein; involved in pre-rRNA processing in the nucleolus; autoregulates splicing of its transcript
YGR105W	<i>VMA21</i>	Verified	7	Integral membrane protein that is required for vacuolar H <sup>+</sup> -ATPase (V-ATPase) function, although not an actual component of the V-ATPase complex; functions in the assembly of the V-ATPase; localized to the yeast endoplasmic reticulum (ER)
YGR118W	<i>RPS23A</i>	Verified	7	Ribosomal protein 28 (rp28) of the small (40S) ribosomal subunit,

					required for translational accuracy; nearly identical to Rps23Bp and similar to E. coli S12 and rat S23 ribosomal proteins; deletion of both RPS23A and RPS23B is lethal
YHR021C	<i>RPS27B</i>	Verified	8		Protein component of the small (40S) ribosomal subunit; nearly identical to Rps27Ap and has similarity to rat S27 ribosomal protein
*YHR022C-A		Unchar.	8		Putative protein of unknown function; identified by gene-trapping, microarray-based expression analysis, and genome-wide homology searching
YHR073C-B		Dubious	8		Dubious open reading frame unlikely to encode a protein, based on available experimental and comparative sequence data; partially overlaps verified ORF YHR073W; identified by expression profiling and mass spectrometry
*YHR073W-A		Dubious	8		Dubious open reading frame unlikely to encode a protein, based on available experimental and comparative sequence data; partially overlaps verified ORF YHR073W; identified by expression profiling and mass spectrometry
YHR141C	<i>RPL42B</i>	Verified	8		Protein component of the large (60S) ribosomal subunit, identical to Rpl42Ap and has similarity to rat L44; required for propagation of the killer toxin-encoding M1 double-stranded RNA satellite of the L-A double-stranded RNA virus
YHR199C-A	<i>NBL1</i>	Verified	8		Subunit of the conserved chromosomal passenger complex (CPC; Ipl1p-Sli15p-Bir1p-Nbl1p), which regulates mitotic chromosome segregation; not required for the kinase activity of the complex; mediates the interaction of Sli15p and Bir1p
*YJL062W-A	<i>COA3</i>	Verified	10		Mitochondrial inner membrane protein that participates in regulation of COX1 translation, Cox1p stabilization, and cytochrome oxidase assembly
*YJR048W	<i>CYCI</i>	Verified	10		Cytochrome c, isoform 1; electron carrier of the mitochondrial intermembrane space that transfers electrons from ubiquinone-cytochrome c oxidoreductase to cytochrome c oxidase during cellular respiration
YJR094W-A	<i>RPL43B</i>	Verified	10		Protein component of the large (60S) ribosomal subunit, identical to Rpl43Ap and has similarity to rat L37a ribosomal protein
*YKR057W	<i>RPS21A</i>	Verified	11		Protein component of the small (40S) ribosomal subunit; nearly identical to Rps21Bp and has similarity to rat S21 ribosomal protein
*YLR061W	<i>RPL22A</i>	Verified	12		Protein component of the large (60S) ribosomal subunit, has similarity to Rpl22Bp and to rat L22 ribosomal protein
*YLR062C	<i>BUD28</i>	Dubious	12		Dubious open reading frame, unlikely to encode a protein; not conserved in closely related <i>Saccharomyces</i> species; 98% of ORF overlaps the verified gene RPL22A; diploid mutant displays a weak budding pattern phenotype in a systematic assay
YLR154C-G		Unchar.	12		Putative protein of unknown function identified by fungal homology comparisons and RT-PCR; this ORF is contained within RDN25-2 and RDN37-2
YLR154W-A		Dubious	12		Dubious open reading frame unlikely to encode a protein; encoded within the the 25S rRNA gene on the opposite strand
YLR154W-B		Dubious	12		Dubious open reading frame unlikely to encode a protein; encoded within the the 25S rRNA gene on the opposite strand
*YLR325C	<i>RPL38</i>	Verified	12		Protein component of the large (60S) ribosomal subunit, has similarity to rat L38 ribosomal protein
*YLR388W	<i>RPS29A</i>	Verified	12		Protein component of the small (40S) ribosomal subunit; nearly identical to Rps29Bp and has similarity to rat S29 and E. coli S14 ribosomal proteins
YML054C-A		Unchar.	13		Putative protein of unknown function

*YML081C-A	<i>ATP18</i>	Verified	13	Subunit of the mitochondrial F1F0 ATP synthase, which is a large enzyme complex required for ATP synthesis; termed subunit I or subunit j; does not correspond to known ATP synthase subunits in other organisms
*YMR272W-A		Dubious	13	Dubious open reading frame unlikely to encode a protein, based on available experimental and comparative sequence data; overlaps the verified gene SCS7/YMR272C
*YOR045W	<i>TOM6</i>	Verified	15	Component of the TOM (translocase of outer membrane) complex responsible for recognition and initial import steps for all mitochondrially directed proteins; promotes assembly and stability of the TOM complex
*YOR302W		Verified	15	CPA1 uORF, Arginine attenuator peptide, regulates translation of the CPA1 mRNA
*YOR312C	<i>RPL20B</i>	Verified	15	Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl20Ap and has similarity to rat L18a ribosomal protein
*YOR335W-A		Dubious	15	Dubious open reading frame unlikely to encode a protein, based on available experimental and comparative sequence data; partially overlaps the verified gene ALA1/YOR335C
*YPR133W-A	<i>TOM5</i>	Verified	16	Component of the TOM (translocase of outer membrane) complex responsible for recognition and initial import of all mitochondrially directed proteins; involved in transfer of precursors from the Tom70p and Tom20p receptors to the Tom40p pore
YPR160C-A		Dubious	16	Identified by gene-trapping, microarray-based expression analysis, and genome-wide homology searching

---



**Table S4.** Top 50 genes emerging from the enrichment analysis of couples and triplets that recapitulate the genome wide analysis of Liti et al.

Genes with high phylogenetic performances	SNP/indel-based analysis on couples		SNP/indel-based analysis on triplets	
	Gene name	Count	Gene name	Count
YAR042W	<b>YNL161W</b>	5018	<b>YML056C</b>	500
YJL099W	<b>YJL057C</b>	4972	<b>YJL057C</b>	408
YJL057C	<b>YML056C</b>	4390	<b>YNL161W</b>	393
YJL051W	<b>YAR042W</b>	4301	YBR139W	358
YKL068W	YBR163W	4101	YML061C	344
YML080W	<b>YML080W</b>	3749	<b>YKL068W</b>	341
YML056C	<b>YJL099W</b>	3592	<b>YAR042W</b>	329
YNL161W	<b>YJL051W</b>	3452	<b>YJL051W</b>	320
YNL125C	<b>YNL125C</b>	3410	YBR132C	314
YOR133W	YBR138C	3364	<b>YML080W</b>	307
YPR152C	YML061C	3363	YBR138C	305
YBL052C	<b>YKL068W</b>	3338	YBR163W	303
	YBL052C	3270	YMR065W	287
	YAL018C	3163	YJR109C	281
	YMR065W	2896	YDR472W	271
	YBR139W	2799	YML059C	269
	YDR472W	2427	YKL121W	264
	YBR132C	2380	<b>YJL099W</b>	263
	YJR109C	2201	<b>YNL125C</b>	258
	YBR133C	1950	YAL018C	251
	<b>YOR133W</b>	1944	YBR136W	248
	YKL121W	1942	YBL052C	243
	YDR046C	1894	YJL056C	239
	YOR132W	1849	YJL085W	239
	YJL050W	1801	YOR132W	231
	YJL092W	1786	YNL152W	229
	YML059C	1776	<b>YOR133W</b>	225
	YBR136W	1684	YJL080C	222
	YNL152W	1683	YDR046C	220
	YMR029C	1672	YJL050W	216
	YJL026W	1665	YJL092W	207
	YJL080C	1640	YML091C	200
	YJL085W	1574	YLR121C	200
	YBR126C	1526	YMR029C	197
	YMR032W	1466	YBR198C	197
	YJL084C	1462	YDR011W	194
	YBR184W	1450	YML023C	193
	YKL062W	1444	YBR133C	188
	YJL054W	1440	YNL138W	186
	YBR198C	1334	YKL062W	185

YMR066W	1318	YOR137C	180
YPL176C	1300	YOR109W	180
YKL022C	1282	YMR066W	178
YDR011W	1273	YBR193C	175
YNL138W	1264	YJL054W	175
YJL056C	1248	YGL151W	172
YNL232W	1244	YNL232W	169
YEL061C	1212	YBR126C	167
YKL046C	1204	YBR184W	165
YOR205C	1177	YHR164C	165

---

**Table S5.** The  $\omega$  ( $dN/dS$ ) ratios for the three different lineages tested, plus the background

Gene name	$dN/dS$				
	average all	West African	Malaysian	Wine/European	Other
YJL099W	0.2059	0.8046	0.0566	0.4006	0.1915
YJL057C	0.2235	1.2880	0.3997	0.0010	0.2500
YJL051W*	0.5575	0.0001	0.8065	999.0	0.4830
YKL068W	0.2347	0.8244	0.2219	30.704	0.2098
YML080W	0.1334	0.1512	0.2270	0.0001	0.0909
YML056C	0.1236	0.1049	0.0001	0.1379	0.1049
YNL161W	--	--	--	--	--
YNL125C	0.1514	--	0.1163	0.0001	0.1576
YOR133W	0.0424	0.8756	0.0001	0.0001	0.0563
YAR042W**	0.1818	0.9973	0.1840	0.5482	0.1620
YBL052C	0.2378	0.0001	999.0	0.2148	0.2315
YBR163W	0.5597	0.5112	89.4128	1.6102	0.4833
YPR152C	0.4786	1.1196	1.3227	0.4387	0.4386

\*LRT is significant when two values of  $\omega$  were evaluated (2,622 for Wine/European cluster, 0,3855 for background)

\*\*LRT is significant when two values if  $\omega$  were evaluated (0,5478 for Wine/European cluster, 0,1640 for background)